Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2007, Article ID 59125, 8 pages
doi:10.1155/2007/59125

Research Article

Secure Hashing of Dynamic Hand Signatures Using Wavelet-Fourier Compression with BioPhasor Mixing and 2^N Discretization

Yip Wai Kuan, Andrew B. J. Teoh, and David C. L. Ngo

Faculty of Information Science and Technology (FIST), Multimedia University, Jalan Ayer Keroh, Bukit Beruang, Melaka 75450, Malaysia

Received 28 February 2006; Revised 23 July 2006; Accepted 18 September 2006

Recommended by Bülent Sankur

We introduce a novel method for secure computation of biometric hash on dynamic hand signatures using BioPhasor mixing and 2^N discretization. The use of BioPhasor as the mixing process provides a one-way transformation that precludes exact recovery of the biometric vector from compromised hashes and stolen tokens. In addition, our user-specific 2^N discretization acts both as an error correction step as well as a real-to-binary space converter. We also propose a new method of extracting compressed representation of dynamic hand signatures using discrete wavelet transform (DWT) and discrete fourier transform (DFT). Without the conventional use of dynamic time warping, the proposed method avoids storage of user's hand signature template. This is an important consideration for protecting the privacy of the biometric owner. Our results show that the proposed method could produce stable and distinguishable bit strings with equal error rates (EERs) of 0% and 9.4% for random and skilled forgeries for stolen token (worst case) scenario, and 0% for both forgeries in the genuine token (optimal) scenario.

1. INTRODUCTION

Recently there is a growing interest in applying biometric for authentication, especially in deriving compact representation of the biometric for cryptographic uses or as encryption keys. This is because existing authentication methods are based on password and cryptographic keys, which are stolen easily, are not sufficiently secure since they are based on what you know and what you have. Using human biometric input provides an additional authentication factor based on what you are. Many good cryptographic techniques exist but since biometric data are not exactly reproducible at each capture, new frameworks and formalisms related to integrating biometrics into cryptosystems need to be considered. In this paper, we chose dynamic hand signatures as the biometric of study for computing biometric keys due to its wide social acceptance and low cost of implementation. We focus on using dynamic hand signatures rather than just off line hand signatures for higher security. Dynamic hand signatures are more difficult to copy as they require the capture of timing information from the signing action and other behavioral characteristics such as the pressure imposed, altitude of the pen and azimuth. Past research into hand signature had focused on increasing the recognition rate between forged and genuine signatures regardless of storage security and capacity. However, in the recent years, with increasing awareness for user privacy especially on the internet and for remote access, there has been a developing body of literature on the protection of biometric data and secure management of keys.

Existing methods of extractions typically require the storage of the template hand signature signals. This is because a template signal is needed to adjust nonlinear variations in the input hand signature signals. A few variations exist for the storage of user templates: (1) use of tamper-proof cards and (2) centralized server. The former which normally requires some form of PIN or password for access permission to the template is not secure as the PIN is meant to be memorized and is hence short and easy to be guessed. Storing the biometric in a centralized server also has serious security implication if the server is compromised.

From the key management perspective, if the biometric template is compromised, the user needs to change both his key and biometric template. Note that for physiological biometric such as fingerprint, iris, DNA, and such, replacement

of biometric secret is not even possible. The solution is to incorporate another independent factor for authentication using random token which could be stored in a tamper-proof card. The biometric would then be combined on the fly with this random token in a one-way manner so that the resulting biometric hash would not reveal information about the biometric. In the event of key compromise, a new key would be reissued using another random token but not the biometric.

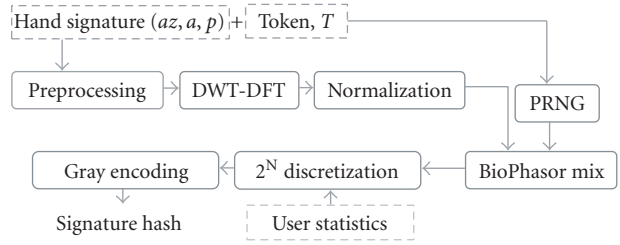

Figure 1: Outline of proposed scheme.

2. LITERATURE REVIEW

There are two main groups of work in this area: (1) pure biometric approach, and (2) biometric + user-token approach. The difference between the two is that the second method incorporates a random token for each user, which provides better security as authentication requires the input of another independent token, which is stored in a tamper-proof device. Another advantage is that the keys are cancelable for the second case as the keys are not solely dependent on the biometric alone. The first biometric hash on dynamic hand signature was proposed by Vielhauer et al. in [1] which used a 50-feature-parameter set from dynamic hand signature and an interval matrix to store the upper and lower thresholds permissible for correct identification. The authors also proposed using written passwords for authentication in [2]. Another scheme similar to [1, 2] is Feng-Chan [3] which also used specific boundaries for each user. The scheme uses 43 features (but not all are published) and reported equal error rate (EER) of 8% but the uniqueness of the output vector is only 1 in 2^{40}. Since these methods are parameter-based, the feature extraction is limited and short, and could not support use in cryptographic systems as they are small in key space, the keys are not cancelable and more importantly, they are generally low in entropy. They are also not secure due to storage of user-specific statistical boundaries that could be used to recover the biometric features. Chang et al. [4] on the other hand used function-based extraction using statistical methods like principal component analysis on face data but the keys are however not cancelable.

Soutar et al. [5], Juels and Wattenberg [6], Juels and Sudan [7], Clancy et al. [8], and Goh and Ngo [9] proposed to incorporate random token into the biometric to allow replaceable or cancelable keys. Soutar et al. [5] proposed the biometric encryption method which required the correlation of the biometric image with a predesigned filter, followed by key generation using a lookup table. Juels and Wattenberg [6] and Juels and Sudan [7], respectively, proposed fuzzy commitment (using error correction codes and XOR) and fuzzy vault which extends the former by using secret sharing to hide the secret. Goh and Ngo [9] used the method of iterative inner product which has the overall effect of random projection of biometric vector based on random token, while preserving the distances within the biometric vector. Yip et al. [10] combined the methods of Goh and Ngo [9] and Chang et al. [4] to enable longer and cancelable or replaceable keys but however, the user-specific key statistics required to correct the feature vector allows an adversary to easily guess the most probable combination from the compromised user boundaries information and reduced number of segments, for example, smaller search space. In another paper [11] Yip et al. proposed a modification to 2^N discretization and fixed the boundaries to deter guessing based on boundaries and segment size. However, the use of iterative inner product of Goh and Ngo introduced a security loophole which enables partial recovery of the biometric feature through multiple extractions of the biometric. In particular, this problem lies in the straightforward multiplication operation of a random basis and the biometric feature which allows an adversary to solve for the biometric vector through QR factorization as shown in [12]. Moreover, the scheme in [11] is globally tuned to provide higher security but the recognition rate is relatively poor.

In this paper, we would like to address the following issues with past implementation of biometric hashing on hand signature.

(i) Hand signature template storage. Many good recognition schemes like Kholmatov-Yanikoglu [13], Feng-Chan [14], Martinez et al. [15], and Hastie et al. [16] used dynamic time warping for distance measure but these methods require storage of the template signature. It is not recommended that the hand signature template be stored in a centralized server or locally in the event of key compromise because the biometric template has permanent association with the user.

(ii) Noncancelable biometric. Previous schemes of Vielhauer et al. [1] and Feng-Chan [3] which relied solely on the biometric are not secure and are inconvenient as the biometric keys cannot be reissued if stolen.

(iii) Poor recognition rate for high security. Previous scheme of Yip et al. [11] assumes globally tuned parameters which do not enhance the local recognition rate, that is, skilled forgery EER is 17% and random forgery EER is 7%.

(iv) Partial leakage of biometric data. The use of iterative inner product [10, 11] leaks partial biometric data if used multiple times.

3. PROPOSED SCHEME

In this paper, we propose a new method for deriving cancelable hand signature key based on the random mixing step of BioPhasor and user-specific 2^N discretization for better recognition rate. We summarize our proposed scheme in Figure. Our proposed scheme utilized dynamic features such as time stamps of signature points, pressure, azimuth, and altitude. A new function-based feature extraction method is introduced, combining discrete wavelet transform (DWT)

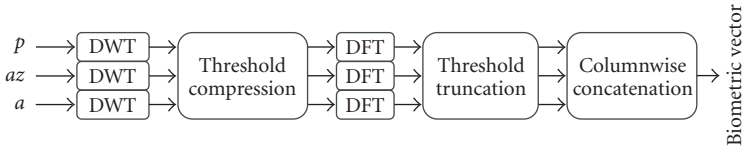

Figure 2: DWT-DFT feature extraction.

for location-sensitive compression and discrete fourier transform (DFT) for frequency analysis to obtain compact representation of the biometric feature. We chose the function-based approach (based on the whole input signals) over parameter-based approach (using specific features such as number of pen ups, duration of signature) to avoid the need to perform individual feature selection and obtain longer keys. The extracted biometric feature is then randomly mixed with a user token T, using the BioPhasor mixing method, which "encrypts" the biometric secret in a one-way manner. A pseudorandom number generator (PRNG) is used to generate a random basis for use in the BioPhasor mixing process. Then, the randomly extracted biometric feature undergoes a real-to-binary conversion using 2^N discretization, with the help of user-specific statistics. Only discretization segment sizes are considered so that they do not reveal original information that could be used for reconstruction of the biometric feature. Finally, we require the use of Gray encoding to ensure that discretized segments that are sufficiently close to the genuine distribution have lower Hamming distances to the genuine index.

3.1. DWT-DFT biometric feature extraction

We applied a hybrid of DWT and DFT methods for the biometric feature extraction as depicted in Figure 2. DWT was chosen due to its good localization feature which gave a more accurate model for compression without losing important information such as sudden peaks or stops as shown by da Silva and de Freitas [17], Deng et al. [18], Lejtman and George [19] and Nakanishi et al. [20]. Then DFT threshold-based compression as by Lam and Kamins [21] was performed on the DWT-compressed vector to further remove high-frequency coefficients, resulting in a very compact representation of the dynamic signature features.

We assume a pressure-sensitive pen and tablet for capturing the online signature signals in terms of pressure information (p), pen altitude (a), and azimuth (az) for each point. Since the points are sampled consistently (10 milliseconds), no resampling was performed and the pen-down segments (detected as points between two pen downs) are concatenated to form one single signal. Each signal is then compressed with the DWT-DFT method described below.

The DWT involves the selection of dyadic (powers of two) scales and positions, and applying them to the mother wavelet. Each dynamic signature signal can be modeled as function $f(t) \in L^2(\mathbb{R})$ that defines space of square integrable functions. Its wavelet transform can be represented as

$$f(t) = \sum_{j=1}^{L} \sum_{k=-\infty}^{\infty} d(j,k)\psi(2^{-j}t - k) + \sum_{k=-\infty}^{\infty} a(L,k)\phi(2^{-L}t - k) \tag{1}$$

with $\psi(t)$ the mother wavelet, $\phi(t)$ the scaling function, and the orthonormal basis for $L^2(\mathbb{R})$ defined by the set of functions

$$\left\{\sqrt{2^{-l}}\phi(2^{-L}t - k), \sqrt{2^{-j}}\psi(2^{-j}t - k) \mid j \le L,\ j,k,L \in \mathbb{Z}\right\}. \tag{2}$$

The approximation coefficients at scale L are defined as

$$a(L,k) = \frac{1}{\sqrt{2^L}} \int_{-\infty}^{\infty} f(t)\phi(2^{-L}t - k)dt \tag{3}$$

while the detail coefficients at scale j are

$$d(j,k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} f(t)\psi(2^{-j}t - k)dt. \tag{4}$$

From (1), the wavelet decomposition at any level L, $f_L(t)$ can be obtained from approximation coefficient $a(L,k)$ and layers of detail coefficients $\{d(j,k) \mid j \le L\}$. The selection of the optimal decomposition level in the hierarchy, however, relies on the experimental data used. In our case, we are also interested in compressing the dynamic hand signature signal by zeroing wavelet coefficients below a certain threshold, to obtain the most compact representation as the feature vector. The global threshold compression method which kept the largest absolute value coefficients and set the others to zero is used. From our experiments (not shown here due to space constraint), we found that the Daubechies 6 (DB6) mother wavelet with decomposition level 2 and compression rate of 60% gave the optimal results, that is, the lowest error rates between genuine and forgery signature distributions.

Each compressed wavelet $F(t) = \text{compress}\,(f_2(t))$ can then be represented by a Fourier integral of form as $g(w) = \int_{-\infty}^{\infty} F(t)e^{-jwt}dt$ with $j = \sqrt{-1}$. The DFT is performed using FFT and the resulting $g(w)$ is then normalized via division by $\sqrt{\sum g_i^2}$ so that $|g| = 1$. Each signal is further truncated using a global thresholding method to obtain $G(w)$. The first 18 significant amplitudes of the transforms are selected based on the lowest error rates (in terms of Euclidean distance) between similarly obtained compressed vectors from genuine

and forgery signatures in experimental database [23]. We used this method to determine the best configuration to ensure that the highest separation between the extracted vectors for the two contrasting classes is retained. In our experiments, we selected the real, imaginary, and magnitude components of (1) tFFT(tDWT(p)) and (2) tFFT(tDWT(a)), and (3) tFFT(tDWT(az)) with tFFT and tDWT being the compression methods described in $F(t)$ and $G(w)$, as dynamic features. Again, these features were chosen because this combination provided the lowest error rates between extracted genuine and forgery vectors from database [23]. Finally, all the DWT-FFT compressed vectors are concatenated and normalized again to form the biometric feature, $b \in \mathbb{R}^n$ of length n. In our experiment, $n = 3$ FFT components $\times 3$ dynamic features $\times 18$ significant amplitudes $= 162$. Note that each original dynamic feature signal is not fixed; each averaged at 208.2 and ranged from 80 to 713 points. Hence, the proposed feature extraction method is able to produce a fixed hash from varying signature inputs.

3.2. Random extraction of biometric feature using BioPhasor

The outline of the BioPhasor [22] mixing step follows.

(1) At enrollment, generate secret random token T using a PRNG, for example, Blum-Blum-Shub generator and store T in a tamper-proof card.

(2) To compute the random basis, generate $m < n$ number of random vectors $t_i \in \mathbb{R}_R^n$ with subscript R denoting that the number is generated randomly using T as the seed, n as the length of the biometric feature, and an integer m. Then, orthonormalize $\forall\, t_i$ using the Gram-Schmidt method.

(3) Compute $h_i = [\sum_{j=1}^{n} \arctan((b_j)^q / t_{i,j})]/n$, where $q \in \mathbb{Z}$ for $i = 1, \ldots, m$. The parameter q tunes the magnitude of the biometric feature element and from our experiment on actual dynamic signature database, $q = 2$ provided the lowest error rate in terms of Euclidean distance between h vectors from genuine and forged signatures.

Since $\arctan(x) + \arctan(x^{-1}) = \pi/2$, the projected vector can be rewritten as $h_i = [\sum_{j=1}^{n}(\pi/2 - \arctan((t_{i,j})/(b_j)^q))]/n$ with $q = 2$, which has a more complicated transformation than random projection using iterative inner product used in earlier work [11]. In particular, the effect is a one-to-one arctan transformation of the random projection of the inverse of biometric vector $b$onto bounded range of $(-\pi/2, \pi/2)$, followed by reflection of the arctan projected space along the x-axis and displacement of $\pi/2$.

3.3. User-specific 2^N discretization

2^N discretization is achieved by dividing the h vector element space into 2^N segments by adjusting to each user standard deviation and the implementation is outlined below.

(1) At enrollment, compute user-specific standard deviation, $\text{std}_i = \sqrt{(\sum_{k=1}^{K}[h_{i,k} - \overline{h_{i,k}}])^2/K}$ for $K = 10$ is the number of training sample, and mean $\overline{h_{i,k}}$, for each element in h.

(2) Estimate and store the number of segments in terms of bit size, n_i for each ith element in h is defined as

$$n_i = \left\{ N \mid \min\left[\text{abs}\left(\frac{\pi/2 - (-\pi/2)}{2^N} - \text{std}_i \times 2 \times \text{k}_{\text{ind}} \right) \right], \; N = 1, \ldots, 30 \right\} \tag{5}$$

for $i = 1, \ldots, n$ and k_{ind}, is a tunable constant for determining how many times of the standard deviation should be considered for the width of each segment. Maximum value of N is arbitrarily limited to 30 to avoid too many bits used for one representative.

(3) At verification, the discretized vector for random projected test input h is d_i, defined as

$$d_i = \left\lfloor \frac{[h_i - (-\pi/2)] \cdot 2^{n_i}}{\pi/2 - (-\pi/2)} \right\rfloor \tag{6}$$

with $\lfloor \cdot \rfloor$ denoting the floor function.

(4) Convert to binary representation using Gray encoding, $p_i = \text{gray}, (d_i)$ because consecutive Gray codes differ by one bit. This ensures that further states from the genuine region, that is, occurring with high probability from imposter test input would have higher Hamming distances.

(5) Perform user-specific permutation by using the PRNG to retrieve another random sequence of indices s (generated based on stored T as the seed), by index sorting to obtain the permutated vector $k_i = p_{s_i}$. The additional permutation step serves to create diffusion to spread the effect of each element to the whole key space.

4. EXPERIMENTS AND SECURITY ANALYSIS

We tested the proposed algorithm with Task 2 training database of the signature verification competition [23], which consists of 40 users with 20 genuine and 20 skilled forgery samples per user. For every user, the first 10 user samples are used as training database to obtain the standard deviations used in our discretization scheme. Since the biometric hashes are in bit strings, we utilized Hamming distance (number of different bits) as the distance measurement. There are two types of forgeries: (1) random and (2) skilled. In random forgery, the forger uses his own signature to access the system. For random forgery error rate evaluation, we compare the remaining 10 test signature keys for each user with every other user. In skilled forgery, the forger simulates the genuine signature that he is impersonating to enter into the system. For skilled forgery, we compare the 20 skilled forgery keys with the 10 test genuine signature keys. To avoid bias in the random numbers, each experiment is repeated 5 times with different random numbers for BioPhasor mixing to obtain the averaged results presented in this section.

Table 1 shows the performance of the various configurations of our proposed algorithm. The Hamming distribution measurements, in terms of mean and standard deviations, are denoted as Mean-[Type] and Std-[Type] with [Type] indicating the distribution for genuine (G), random forgery

Table 1: Comparison of EER and mean distribution for various configurations of k_{ind}.

Type	k_{ind}	Bits	EER-R	EER-S	Mean-G	Std-G	Mean-R	Std-R	Mean-S	Std-S
Bio−Euclidean	2	162	10.833	11.780	0.014	0.008	0.378	0.015	0.149	0.016
Bio−discretized	3.0	1610	0.000	14.278	0.017	0.008	0.447	0.011	0.048	0.018
Bio+gen token	0.2	1618	0.000	0.000	0.061	0.030	0.253	0.025	0.246	0.008
Bio+gen token	0.4	1607	0.000	0.000	0.061	0.030	0.346	0.075	0.246	0.008
Bio+gen token	0.6	1560	0.000	0.000	0.131	0.030	0.449	0.012	0.279	0.008
Bio+gen token	0.8	1510	0.000	0.000	0.101	0.031	0.441	0.013	0.261	0.009
Bio+gen token	1	1470	0.000	0.000	0.083	0.031	0.435	0.013	0.249	0.010
Bio+gen token	2	1310	0.000	0.000	0.069	0.029	0.430	0.013	0.239	0.010
Bio+gen token	3	1209	0.000	0.000	0.060	0.027	0.426	0.013	0.232	0.011
Bio+gen token	4	1148	0.000	0.000	0.035	0.018	0.410	0.014	0.205	0.012
Bio+gen token	5	1088	0.000	0.000	0.025	0.014	0.399	0.014	0.186	0.013
Bio+gen token	6	1047	0.000	0.000	0.020	0.011	0.390	0.015	0.171	0.014
Bio+gen token	7	1007	0.000	0.000	0.017	0.010	0.383	0.015	0.159	0.015
Bio+stolen token	0.2	1620	10.893	10.139	0.061	0.030	0.151	0.035	0.144	0.028
Bio+stolen token	0.4	1606	5.968	10.250	0.061	0.030	0.279	0.121	0.143	0.028
Bio+stolen token	0.6	1563	0.752	10.056	0.059	0.029	0.392	0.076	0.141	0.028
Bio+stolen token	0.8	1503	0.051	9.750	0.055	0.027	0.420	0.043	0.138	0.028
Bio+stolen token	**1.0**	**1468**	**0.000**	**9.389**	**0.051**	**0.026**	**0.425**	**0.032**	**0.133**	**0.028**
Bio+stolen token	2.0	1312	0.000	9.750	0.036	0.019	0.425	0.021	0.107	0.029
Bio+stolen token	3.0	1208	0.000	10.056	0.027	0.015	0.416	0.021	0.086	0.028
Bio+stolen token	4.0	1151	0.000	10.444	0.021	0.012	0.409	0.021	0.071	0.025
Bio+stolen token	5.0	1091	0.000	11.333	0.018	0.010	0.403	0.021	0.061	0.023
Bio+stolen token	6.0	1046	0.000	11.833	0.015	0.009	0.401	0.021	0.053	0.021
Bio+stolen token	7.0	1012	0.000	12.111	0.014	0.008	0.398	0.022	0.047	0.019

(R), and skilled forgery (S) keys generated from our proposed algorithm. When only biometric features are extracted using the proposed DWT-FFT and measure in Euclidean distances, the random forgery equal error rate (EER-R) is 10.8% while the skilled forgery equal error rate (EER-S) is 11.7%. Using 2^N discretization on biometric (bio-discretized) yields better results than using the Euclidean distance alone for the random forgery scenario. However, for the skilled forgery case, the discretization deteriorates the results from 11.7% to 14.2%. This shows that using biometric alone as the key is not sufficient as the entropy is still low and hence, insufficient for providing good separation between genuine and skilled forgery cases.

When the biometric is combined with genuine random token (Bio+gen token), perfect separation is observed for both types of forgeries, that is, random forgery with imposter own token EER-R is 0% and skilled forgery with imposter own token EER-S is 0% (Figure 3). However, this former scenario is only applicable for the case where the user never loses his token which is not realistic in the real world. We simulate the worst case for stolen token scenario (Bio+stolen token) by using the same set of random numbers on all the users. Figure 4 shows the optimal configuration (assuming the stolen token scenario) when $k_{ind} = 1$ which provided EER-R is 0% and EER-S is 9.39%. Figures 5 and 6 illustrate the Hamming distribution for worst and optimal cases. Note also that the mean distribution of the random forgery bit strings in both cases peaks around 0.45, indicating that each user bit strings differ by 45% which is desirable.

We consider possible ways the proposed scheme may be attacked and discuss how the scheme circumvents these attacks.

(1) Brute force attack

The adversary does not have any knowledge of the user key, token or key statistics. He performs a brute force attack by trying out all possible combinations of the key. The computational complexity to guess the key is $2^{n'}$ where n' is the key length.

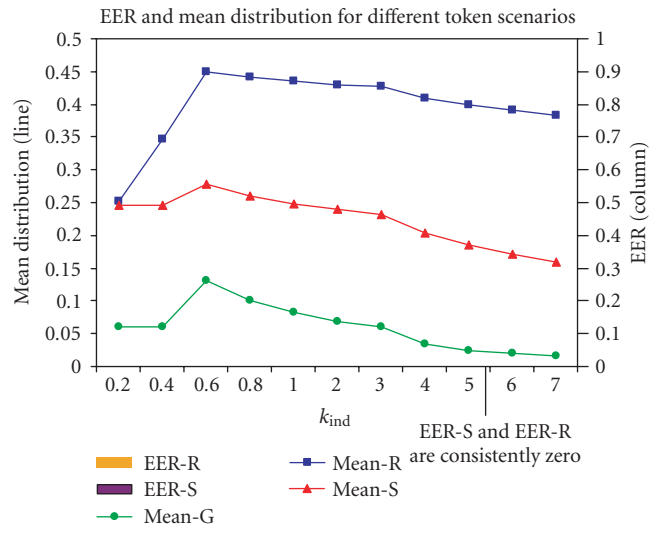

Figure 3: EER and mean distribution for genuine token scenario.

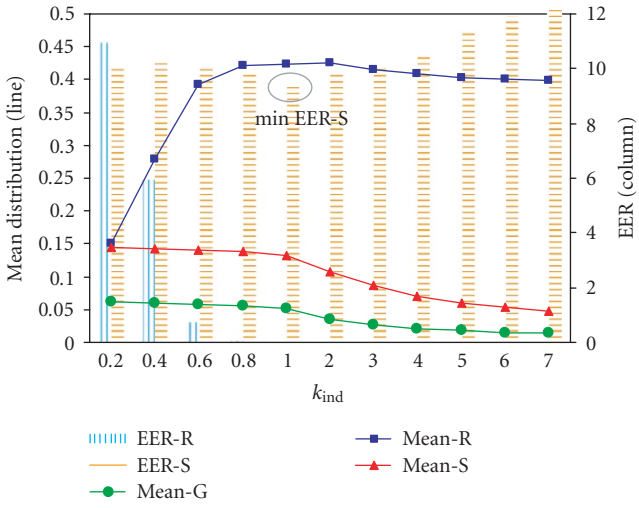

Figure 4: EER and mean distribution for stolen token scenario.

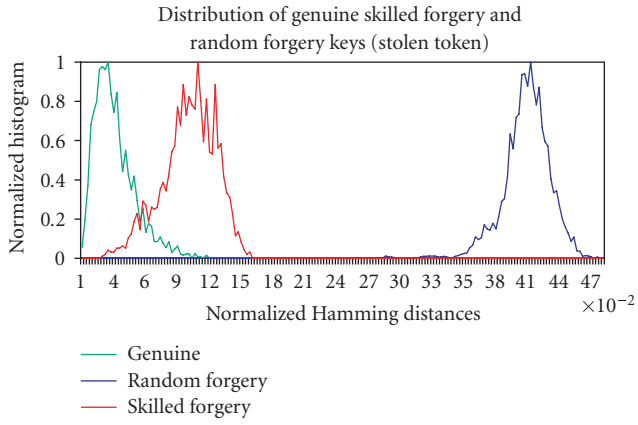

Figure 5: Hamming distribution for stolen token scenario $k_{ind} = 1$.

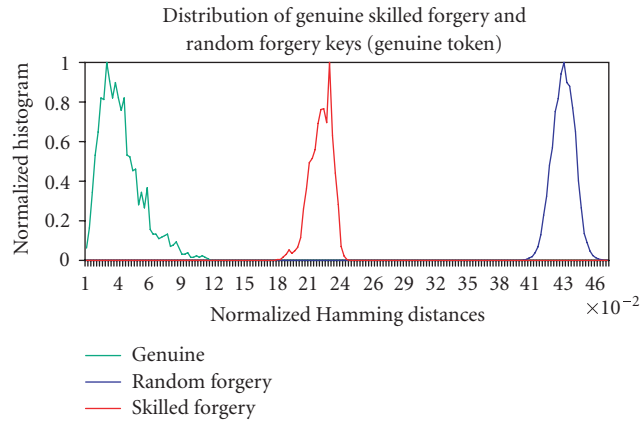

Figure 6: Hamming distribution for genuine token scenario k_{ind}=1.

(2) Multiple key attacks

The adversary eavesdrops on a genuine user to collect multiple keys. Since the genuine token is not compromised, the irreversible random extraction method used in the scheme would ensure that recovery of the biometric feature is performed at least nonpolynomial time (shown in Proposition 3 later in this section).

(3) Substitution token with skilled forgery attack

In this scenario, the imposter uses his own token and skilled forgeries of the genuine signature to hopefully generate a false acceptance into the system. The experimental results are shown in Table 1 as skilled forgery under Bio+genuine token category, with EER-S $\sim$ 0% indicating that using the proposed scheme, this attack will not be successful.

(4) Known key, token, and user statistics attack

This represents the worst case scenario whereby the adversary has access to the key statistics, that is, segment size information and genuine user token. He attempts to combine the stolen token with forged signature to enter the system. From Table 1, for the optimal case where $k_{ind}=1$ is used, the probability of success is with this type of forged entry $\sim$9.39%.

The security strength of the proposed scheme lies in two important dispositions: (1) irreversible random extraction of biometric information via BioPhasor, and (2) transformation from real-valued biometric feature to index space and finally to binary bit strings, which can be seen as a form of error correction to compensate for noisy biometric data as well as lossy compression. The overall effect of these two steps is a one-way transformation of the real-space biometric vector into binary-space hashes without compromising the biometric secret itself. We proceed to show the security proofs of the proposed scheme.

Proposition 1. *If the BioPhasor vector h and the genuine token T are known, recovering the biometric feature b exactly cannot be performed in polynomial time, that is, intractable problem.*

Proof. The random vectors t_i are known since token T is known. Hence, we can form the system of equations $h_i = [\sum_{j=1}^{n} \arctan((b_j)^q/t_{i,j})]/n = [\sum_{j=1}^{n}(\pi/2 - \arctan(t_{i,j}/(b_j)^q))]/n$ where $q \in \mathbb{Z}$ and $i = 1,\ldots,m < n$. Due to the presence of arctan operation, solving the system of equations cannot be straightforwardly performed using QR factorization. Also, since $\tan(h_i) \neq \sum_{j=1}^{n}(b_j)^q t_{i,j}^{-1}$, we cannot linearly transform the system of equations into Gaussian eliminated form. Using Taylor series representation, $\arctan(t_{i,j} \cdot b_j^{-q}) = \sum_{a=0}^{\infty}((-1)(t_{i,j} \cdot b_j^{-q})^{2a+1}/(2a+1))$, we can rewrite the system of equations as $\sum_{j=1}^{n}[\sum_{a=0}^{K}(\pi/2 - (-1)(t_{i,j} \cdot b_j^{-q})^{2a+1}/(2a+1))] = h_i$ assuming that we truncate the series to $K > 1$ terms for approximation purpose. It is clear that this system cannot be solved in polynomial time, hence solving for the biometric b is an intractable problem. □

Proposition 2. *2^N discretization is an irreversible process.*

Proof. Let the 2^N discretization be defined as $f \circ g$ where $f : (-\pi/2, \pi/2)^n \to \mathbb{Z}_{2^N}^n$ and $g : \mathbb{Z}_{2^N}^n \to \{0,1\}^{n'}$ with $n' > n$. Since f is a transformation from real-to-index space, information will be lost. In particular, the continuous to discrete entropy lost is $\log(2^{n_i})$ based on individual segment size n_i as mentioned in Cover and Thomas [24]. Hence the 2^N discretization is irreversible. □

Proposition 3. *The sequence of BioPhasor mixing and 2^N discretization obeys the product principle, that is, the proposed scheme is a one-way transformation.*

Lemma 1. *The product principle of Shannon [25] states that the systematic cascading of different types of ciphers in single cryptosystems will increase the cipher strength provided that the product ciphers are associative but not commutative.*

Proof. Let individual BioPhasor mixing be defined as $f_i : \mathbb{R}^n \times \mathbb{R}^n \to (-\pi/2, \pi/2)$ and let the 2^N discretization be $g : (-\pi/2, \pi/2)^n \to \{0,1\}^{n'}$ with $n' > n$. Clearly $\{f_i\}_{i=1}^{n} \circ g$ is associative but not commutative since the domain and range cannot be interchanged. Since f and g are irreversible from Propositions 1 and 2, and due to the product principle (Lemma 1), $\{f_i\}_{i=1}^{n} \circ g$ is a one-way transformation. □

5. CONCLUSION

We believe that the proposed integration of BioPhasor random mixing and user-specific 2^N discretization is a secure method for deriving biometric hashes from dynamic hand signatures without jeopardizing the privacy of the user. Firstly, the avoidance of using dynamic time warping technique removes the necessity of signature template storage. Secondly, the random token is an independent factor from the biometric, hence if compromised, a new token and hash can be generated but the biometric could still be retained. A stronger notion of security in BioPhasor compared to successive inner product mixing used in our earlier scheme also prevents leakage of sensitive biometric information. We achieve this via the use of arctan for limiting the range as well as for disabling recovery of biometric vector using QR factorization. Meanwhile, the application of user-specific 2^N discretization and DWT-DFT feature extraction enabled better recognition rate especially for the stolen token scenario with EER-R $\sim$ 0% and EER-S $\sim$ 9.4%. By imposing the number of segments to 2^N and limiting the boundaries across the population, we force the adversary to attempt all combinations of the segment space, hence, prevent guessing based on user space.

REFERENCES

[1] C. Vielhauer, R. Steinmetz, and A. Mayerhorf, "Biometric hash based on statistical features of online signatures," in *Proceedings of 16th International Conference on Pattern Recognition (ICPR '02)*, vol. 1, pp. 123–126, Quebec, Canada, August 2002.

[2] C. Vielhauer and R. Steinmetz, "Handwriting: feature correlation analysis for biometric hashes," *EURASIP Journal on Applied Signal Processing*, vol. 2004, no. 4, pp. 542–558, 2004, special issue on Biometric Signal Processing.

[3] H. Feng and C. W. Chan, "Private key generation from on-line handwritten signatures," *Information Management & Computer Security*, vol. 10, no. 4, pp. 159–164, 2002.

[4] Y.-J. Chang, W. Zhang, and T. Chen, "Biometrics-based cryptographic key generation," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '04)*, vol. 3, pp. 2203–2206, Taipei, Taiwan, June 2004.

[5] C. Soutar, D. Roberge, A. Stoianov, R. Gilroy, and B. V. Kumar, "Biometric encryption using image processing," in *Optical Security and Counterfeit Deterrence Techniques II*, vol. 3314 of *Proceedings of SPIE*, pp. 178–188, San Jose, Calif, USA, January 1998.

[6] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in *Proceedings of the 6th ACM Conference on Computer and Communications Security (CCS '99)*, G. Tsudik, Ed., pp. 28–36, Singapore, November 1999.

[7] A. Juels and M. Sudan, "A fuzzy vault scheme," in *Proceedings of IEEE International Symposium on Information Theory (ISIT '02)*, A. Lapidoth and E. Teletar, Eds., p. 408, Lausanne, Switzerland, June-July 2002.

[8] T. C. Clancy, N. Kiyavash, and D. J. Lin, "Secure smartcard-based fingerprint authentication," in *Proceedings of the ACM SIGMM Workshop on Multimedia Biometrics Methods and Applications (WBMA '03)*, pp. 45–52, Berkley, Calif, USA, November 2003.

[9] A. Goh and D.-C. L. Ngo, "Computation of cryptographic keys from face biometrics," in *Proceedings of 7th IFIP International Conference on Communications and Multimedia Security (CMS '03)*, vol. 2828 of *Lecture Notes in Computer Science*, pp. 1–13, Torino, Italy, October 2003.

[10] W.-K. Yip, A. Goh, D.-C. L. Ngo, and A.-B. J. Teoh, "Generation of replaceable cryptographic keys from dynamic handwritten signatures," in *Proceedings of International Conference on Advances in Biometrics (ICB '06)*, vol. 3832 of *Lecture Notes in Computer Science*, pp. 509–515, Hong Kong, 2006.

[11] W.-K. Yip, A. Goh, D.-C. L. Ngo, and A.-B. J. Teoh, "Cryptographic keys from dynamic hand-signatures with biometric secrecy preservation and replaceability," in *Proceedings of the 4th IEEE on Automatic Identification Advanced Technologies (AutoID '05)*, pp. 27–32, Buffalo, NY, USA, October 2005.

[12] K. Liu, H. Kargupta, and J. Ryan, "Random projection-based multiplicative data perturbation for privacy preserving distributed data mining," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 1, pp. 92–106, 2006.

[13] A. Kholmatov and B. Yanikoglu, "Biometric authentication using online signatures," in *Proceedings of the 19th International Symposium on Computer and Information Sciences (ISCIS '04)*, vol. 3280 of *Lecture Notes in Computer Science*, pp. 373–380, Kemer-Antalya, Turkey, October 2004.

[14] H. Feng and C. W. Chan, "Online signature verification using a new extreme points warping technique," *Pattern Recognition Letters*, vol. 24, no. 16, pp. 2943–2951, 2003.

[15] J. C. R. Martinez, J. J. V. Lopez, and F. J. L. Rosas, "A low-cost system for signature recognition," in *Proceedings of International Congress on Research in Electrical and Electronics Engineering (ELECTRO '02)*, 2002.

[16] T. Hastie, E. Kishon, M. Clark, and J. Fan, "A model for signature verification," in *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, vol. 1, pp. 191–196, Charlottesville, VA, USA, October 1991.

[17] A. Vergara da Silva and D. Santana de Freitas, "Wavelet-based compared to function-based on-line signature verification," in *Proceedings of the 15th Brazilian Symposium on Computer Diagramics and Image Processing (SIBGRAPI '02)*, pp. 218–225, Fortaleza-CE, Brazil, October 2002.

[18] P. S. Deng, H.-Y.M. Liao, C. W. Ho, and H.-R. Tyan, "Wavelet-based off-line handwritten signature verification," *Computer Vision and Image Understanding*, vol. 76, no. 3, pp. 173–190, 1999.

[19] D. Z. Lejtman and S. E. George, "On-line handwritten signature verification using wavelets and back-propagation neural networks," in *Proceedings of 6th International Conference on Document Analysis and Recognition (ICDAR '01)*, pp. 992–996, Seattle, Wash, USA, September 2001.

[20] I. Nakanishi, N. Nishiguchi, Y. Itoh, and Y. Fukui, "On-line signature verification method utilizing feature extraction based on DWT," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS '03)*, vol. 4, pp. IV-73–IV-76, Bangkok, Thailand, May 2003.

[21] C. F. Lam and D. Kamins, "Signature recognition through spectral analysis," *Pattern Recognition*, vol. 22, no. 1, pp. 39–44, 1989.

[22] A.-B. J. Teoh and D.-C. L. Ngo, "Cancellable biometrics realization through biophasoring," in *Proceedings of 9th IEEE International Conference on Control, Automation, Robotics and Vision (ICARCV '06)*, Singapore, December 2006.

[23] SVC, First International Signature Verification Competition, 2004, http://www.cs.ust.hk/svc2004/.

[24] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1991.

[25] C. E. Shannon, "Communication Theory of Secrecy Systems," *Bell Systems Technical Journal*, vol. 28, pp. 656–715, 1949.

Yip Wai Kuan received her B.S. and M.S. degrees in computer science from University Science of Malaysia in 1999 and 2003, respectively. Currently, she works as an Analytic Solutions Development Engineer in Intel Malaysia and is completing her Ph.D. degree in information technology from Multimedia University, Malaysia. Her research interests cover the areas of dynamic hand signatures, signal processing, and information security.

Andrew B. J. Teoh received his B.Eng. degree (electronics) in 1999 and Ph.D. degree in 2003 from the National University of Malaysia. He is currently a Senior Lecturer and Associate Dean of the Faculty of Information Science and Technology, Multimedia University, Malaysia. He held the post of cochair (Biometrics Division) in the Center of Excellence in Biometrics and Bioinformatics at the same university. He also serves as a Research Consultant for Corentix Technologies in the research of biometrics system development and deployment. His research interests are in multimodal biometrics, pattern recognition, multimedia signal processing, and internet security. He has published more than 90 international refereed journals and conference papers.

David C. L. Ngo is an Associate Professor and the Dean of the Faculty of Information Science & Technology at Multimedia University, Malaysia since 1999. He was awarded a BAI in microelectronics and electrical engineering and Ph.D. degree in computer science in 1990 and 1995, respectively, both from Trinity College Dublin. His research lies in the area of automation screen design, aesthetic systems, biometric encryption, and knowledge management. He is also author and coauthor of over 20 invited and refereed papers. He is a Member of Review Committee of Displays and Multimedia Cyberspace.